PROTECT Yourself

■ อุรุพจน์ กัลยาสิริ

ตรวจจับผู้บุกรุก

: ปัญหาของตัวตรวจจับการบุกรุก

อธิบายถึงปัญหาสำคัญ ที่ทำให้ระบบตรวจจับการบุกรุก ทำงานไม่ได้ดีเท่าที่ควร

เมื่อทราบถึงความสำคัญของการรักษาความปลอดภัยให้แก่ระบบคอมพิวเตอร์แล้ว ผู้ดูแลระบบมีความจำเป็นอย่างยิ่งที่จะต้องเตรียมการให้พร้อมสำหรับรองรับการเข้ามาของผู้ไม่ประสงค์ดี ซึ่งวิธีการจัดการกับผู้บุกรุกนั้นทำได้หลายวิธี วิธีหนึ่งที่นิยมกันมากก็คือการนำตัวตรวจจับการบุกรุกอัตโนมัติมาใช้งาน โดยในความเป็นจริงแล้วตัวตรวจจับการบุกรุกนั้นมิใช่เครื่องมือที่มีความสามารถรอบด้าน หรือสามารถที่จะจัดการกับผู้ไม่ประสงค์ดีได้ทุกรูปแบบ หากแต่เป็นเครื่องมือชนิดหนึ่งที่เป็นส่วนช่วยในการสร้างความมั่นใจให้แก่ระบบคอมพิวเตอร์ในระดับหนึ่งเท่านั้น

ในปัจจุบันมีงานวิจัยและการคิดค้นจำนวนมากที่กำลังมองหาความเป็นไปได้ที่จะพัฒนาตัวตรวจจับการบุกรุกให้มีประสิทธิภาพมากยิ่งขึ้นโดยนำเทคนิคต่าง ๆ มาช่วยในการออกแบบและพัฒนาเทคนิคและวิธีการต่าง ๆ เหล่านี้ได้แก่ การนำเทคนิคทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) หรือดาต้าไมนิง (Data mining) มาช่วยในการวิเคราะห์เหตุการณ์เพื่อตัดสินใจว่าเหตุการณ์ที่ระบบคอมพิวเตอร์กำลังประสบอยู่เป็นเหตุการณ์ปกติหรือเป็นการบุกรุก หรือการนำเทคนิคเหล่านี้มาแบ่งประเภทของข้อมูลที่น่าสงสัยและกำจัดข้อมูลที่ไม่เกี่ยวข้องกับการวิเคราะห์การบุกรุก การออกแบบโปรแกรมการทำงานให้สามารถทำงานได้อย่างรวดเร็วเป็นแบบเวลาจริง (Real-time) และสิ้นเปลืองทรัพยากรน้อยที่สุดก็เป็นหัวข้อที่กำลังเป็นที่สนใจกันมากในปัจจุบัน ถึงแม้ว่าจะมีงานวิจัยจำนวนมากที่กำลังศึกษากันอยู่ แต่ก็ยังไม่สามารถกำจัดปัญหาสำคัญของตัวตรวจจับการบุกรุกให้หมดไปได้อย่างสิ้นเชิง ปัญหาสำคัญเหล่านี้ได้แก่ ***การจัดการกับข้อมูลจำนวนมาก การวิเคราะห์และตัดสินใจที่มีประสิทธิภาพ และการทำงานให้ได้รวดเร็วและทันท่วงที***

ต้องจัดการกับข้อมูลจำนวนมาก

เนื่องจากตัวตรวจจับอัตโนมัติจะต้องวิเคราะห์เหตุการณ์เพื่อสรุปว่ามีการบุกรุกเข้ามาทำลายระบบคอมพิวเตอร์หรือไม่ ซึ่งจะต้องอาศัยการวิเคราะห์ข้อมูลจำนวนมาก โดยข้อมูลเหล่านี้จะแตกต่างกันตามประเภทของตัวตรวจจับนั้น ๆ เช่น ถ้าแบ่งประเภทของตัวตรวจจับตามลักษณะของข้อมูลที่นำมาวิเคราะห์จะสามารถจัดได้เป็นสองประเภทคือ ตัวตรวจจับเฉพาะตน (Host-based Intrusion Detection System) และตัวตรวจจับทุกโฮสต์ในเครือข่าย (Network-based Intrusion Detection System)

ตัวตรวจจับเฉพาะตนจะรับข้อมูลที่ส่งถึงตนเอง หรือเป็นข้อมูลที่มีความเกี่ยวข้องโดยตรงกับระบบของตนเอง เช่น แพ็กเกตข้อมูลที่ส่งมายังระบบ หรือแฟ้มบันทึกการทำงานต่าง ๆ หากพิจารณาเครือข่ายคอมพิวเตอร์เฉพาะบริเวณหรือเครือข่ายแลนนั้นจะเห็นได้ชัดเจนว่าแพ็กเกตข้อมูลที่ส่งมายังเครื่องคอมพิวเตอร์เครื่องใดเครื่องหนึ่งจะมีปริมาณมากน้อยจะขึ้นอยู่กับการใช้งานของระบบในขณะนั้น เช่น หากระบบกำลังให้บริการถ่ายโอนแฟ้มข้อมูลหรือ เอฟทีพี (FTP) ปริมาณข้อมูลอาจมีจำนวนมากและรับส่งกันในอัตราที่สูง หรือหากระบบเป็นเซิร์ฟเวอร์ที่ให้บริการเว็บก็จะมีแพ็กเกตที่รับส่งจากเครือข่ายภายนอกจำนวนมากหรือน้อยขึ้นอยู่กับปริมาณการเข้าชมเว็บนั้น ลักษณะของแพ็กเกตข้อมูลที่ตัวตรวจจับจะนำมาพิจารณานั้นจะต้องผ่านการกรองเสียก่อน เนื่องจากว่าข้อมูลทั้งหมดมีปริมาณมหาศาลและก็มีข้อมูลจำนวนมากอีกเช่นกันที่ไม่มีประโยชน์ใด ๆ ต่อการวิเคราะห์การบุกรุก ตัวตรวจจับการบุกรุกจะใช้เทคนิคในการกรองข้อมูล เลือกวิเคราะห์เฉพาะข้อมูลที่คาดว่าจะมีความสัมพันธ์กับการบุกรุกในขณะนั้น สำหรับแพ็กเกตข้อมูลในเครือข่ายแล้วตัวตรวจจับการบุกรุกมักจะเลือกดูเฉพาะเฮดเดอร์ (Header) หรือส่วนหัวของแพ็กเกตที่ใช้ในการควบคุมการส่งผ่านข้อมูลเท่านั้น เพราะหากจะต้องวิเคราะห์แพ็กเกตทั้งหมดแล้วอาจจะไม่มีประโยชน์และสิ้นเปลือง

ทรัพยากรของระบบไป ลักษณะของแพ็กเกตข้อมูลในเครือข่ายจะต้องประกอบไปด้วยสองส่วนสำคัญคือ ส่วนเฮดเดอร์ที่ใช้ในกระบวนการทำงานของโปรโตคอล และส่วนของข้อมูลที่ต้องการส่ง ดังแสดงในรูปที่ 1

เฮดเดอร์	ข้อมูล

รูปที่ 1
แพ็กเกตข้อมูลในเครือข่าย

โดยปกติการวิเคราะห์การบุกรุกจากเฮดเดอร์ของแพ็กเกตข้อมูลมักจะสังเกตจากพฤติกรรมของแพ็กเกตนั้น ๆ ว่าผิดปกติหรือไม่และเป็นไปตามขั้นตอนการทำงานของโปรโตคอลที่กำหนดไว้หรือไม่ หากไม่เป็นไปตามขั้นตอนของโปรโตคอลปกติก็อาจเป็นแพ็กเกตที่น่าสงสัยได้ แต่ก็มิใช่จะสรุปได้ทันทีว่าจะเป็นการบุกรุกที่เกิดขึ้นจริง เพราะจำเป็นต้องอาศัยข้อมูลอื่นรอบข้างช่วยในการพิจารณาด้วยเช่นกัน

```
host.domain - - [01/Nov/2000:04:33:05 +0700] "GET /a.html HTTP/1.0" 200 740
host.domain - - [01/Nov/2000:04:34:04 +0700] "HEAD /b.html HTTP/1.1" 200 0
host.domain - - [01/Nov/2000:04:37:10 +0700] "HEAD /c.html HTTP/1.1" 200 0
host.domain - - [01/Nov/2000:04:38:47 +0700] "GET /d.html HTTP/1.0" 301 286
host.domain - - [01/Nov/2000:04:38:52 +0700] "GET /PAPER HTTP/1.0" 301 281
host.domain - - [01/Nov/2000:04:39:57 +0700] "GET /mesg HTTP/1.0" 200 48941
host.domain - - [01/Nov/2000:04:40:26 +0700] "HEAD /03.htm HTTP/1.1" 200 0
host.domain - - [01/Nov/2000:04:40:26 +0700] "GET /4.htm HTTP/1.0" 200 43630
host.domain - - [01/Nov/2000:04:41:11 +0700] "GET /s.html HTTP/1.0" 200 6578
...                ...                ...                ...
```

รูปที่ 2 แฟ้มบันทึกการเรียกดูเว็บเพจของเว็บเซิร์ฟเวอร์

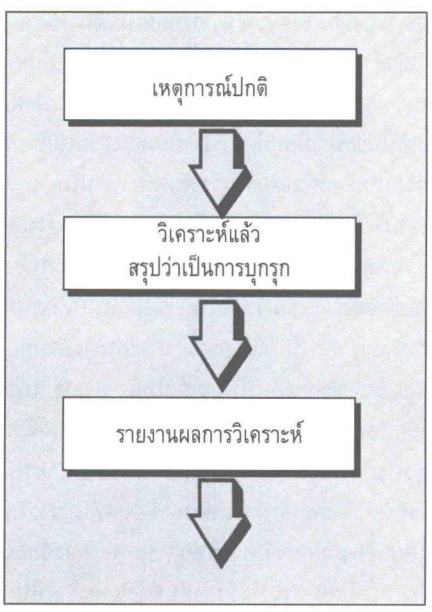
นอกจากแพ็กเกตข้อมูลแล้ว ระบบคอมพิวเตอร์ก็ยังมีแฟ้มบันทึกการทำงานต่าง ๆ เก็บไว้ด้วยเช่นกัน แฟ้มบันทึกการทำงานเหล่านี้จะเก็บเหตุการณ์หรือสิ่งต่าง ๆ ที่เกิดขึ้นกับระบบว่ามีอะไรเกิดขึ้นบ้างทั้งกับระบบปฏิบัติการรวมถึงซอฟต์แวร์ที่ทำงานอยู่ หากพิจารณาระบบปฏิบัติการยูนิกซ์แล้วจะเห็นได้ชัดเจนว่าระบบจะเก็บข้อมูลเหตุการณ์ต่าง ๆ ที่เกิดขึ้นในรูปของแฟ้มบันทึกการทำงานของระบบตั้งแต่เริ่มการทำงานระบบจนถึงการหยุดการทำงาน สิ่งต่าง ๆ ที่เกิดขึ้นจะถูกบันทึกเก็บไว้ในแฟ้มพิเศษที่ผู้ดูแลระบบจะนำมาวิเคราะห์ดูว่ามีสิ่งใดผิดปกติ และเกิดจากสาเหตุอะไร นอกจากนั้นระบบคอมพิวเตอร์ที่ให้บริการใดก็มักจะมีแฟ้มบันทึกการทำงานสำหรับการให้บริการนั้นด้วยเช่นกัน ดังเช่นตัวอย่างในรูปที่ 2 เป็นแฟ้มบันทึกการทำงานการเข้ามาดูเว็บของระบบ

ตัวตรวจจับการบุกรุกอีกประเภทได้แก่ ตัวตรวจจับการบุกรุกทุกโฮสต์ในเครือข่าย จะทำงานต่างจากแบบแรกคือ จะพิจารณาแพ็กเกตข้อมูลทั้งหมดที่อยู่ในเครือข่ายโดยไม่คำนึงว่าแพ็กเกตข้อมูลนั้นจะส่งถึงโฮสต์ใด เนื่องจากว่าจำเป็นจะต้องวิเคราะห์ว่ามีโฮสต์ใดบ้างในเครือข่ายถูกบุกรุก ดังนั้นข้อมูลที่อยู่ในเครือข่ายทั้งหมดจะต้องนำมาวิเคราะห์ ข้อมูลทั้งหมดนี้มีจำนวนมากหากเปรียบเทียบกับการทำงานของตัวตรวจจับเฉพาะตน ดังนั้นจึงมีความจำเป็นอย่างยิ่งที่จะต้องนำเทคนิคต่าง ๆ มาใช้ในการแบ่งแยกข้อมูลที่จะวิเคราะห์กับข้อมูลที่ไม่มีความสำคัญ

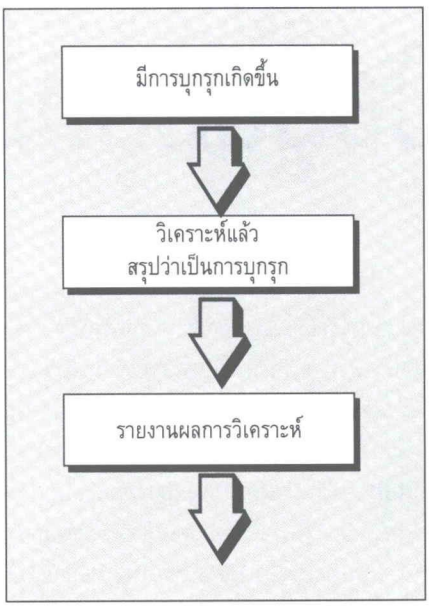
จะเห็นได้อย่างชัดเจนว่าตัวตรวจจับการบุกรุกจะต้องจัดการกับข้อมูลจำนวนมากและหลากหลายชนิด ดังนั้นจึงจำเป็นจะต้องมีวิธีการเลือกวิเคราะห์และพิจารณาให้เหมาะสมด้วย วิธีการที่มักนำมาใช้ในการจัดการกับข้อมูลจำนวนมาก ได้แก่ การแบ่งประเภทข้อมูล (Data Clustering) ซึ่งเป็นเทคนิคในการแบ่งว่าข้อมูลประเภทใดจะนำมาพิจารณา หรือข้อมูลประเภทใดไม่มีความสำคัญ เทคนิคอีกอย่างคือ การเลือกดูส่วนสำคัญ (Feature Selection) เป็นการเลือกดูเฉพาะจุดที่มีความสำคัญต่อการวิเคราะห์การบุกรุกเท่านั้น เช่น เลือกดูเฉพาะแพ็กเกตที่มีเฮดเดอร์บางแบบ เป็นต้น

การสรุปผลการวิเคราะห์ที่ถูกต้อง

เมื่อตัวตรวจจับการบุกรุกได้วิเคราะห์ข้อมูลแล้วจำเป็นจะต้องสรุปให้ผู้ดูแลระบบทราบว่าเหตุการณ์ที่ได้วิเคราะห์ไปนั้น ถือเป็นสิ่งปกติหรือผิดปกติอย่างไร หากตัวตรวจจับการบุกรุกอัตโนมัติสามารถสรุปเหตุการณ์ที่เกิดขึ้นได้อย่างถูกต้องก็จะเป็นสิ่งที่ดีและมีประโยชน์ แต่ถ้าสรุปแล้วผิดจากความเป็นจริงไปมากก็อาจจะส่งผลเสียต่อการทำงานได้เช่นเดียวกัน ดังนั้นการสรุปผลการวิเคราะห์จำเป็นต้องมีความระมัดระวังเป็นอย่างยิ่ง เทคนิคที่นำมาใช้ในการวิเคราะห์เหตุการณ์ ได้แก่ การวิเคราะห์ทางสถิติ การวิเคราะห์โดยใช้การเปรียบเทียบพฤติกรรม การวิเคราะห์โดยใช้กฎ เป็นต้น



รูปที่ 3 การตัดสินใจผิดทางบวก



➲ *รูปที่ 4*
การตัดสินใจผิดทางลบ

หากมีการสรุปผลการวิเคราะห์ที่ผิดพลาดอาจเกิดความเสียหายได้ รูปแบบของการสรุปผลที่ผิดพลาดมีสองชนิด ชนิดแรกคือ การตัดสินใจผิดทางบวก (False positive) อีกชนิดคือ การตัดสินใจผิดทางลบ (False negative) การตัดสินใจผิดทางบวกเป็นการทำงานที่ผิดพลาดของตัวตรวจจับการบุกรุกที่มักพบได้ง่ายที่สุด อาจเป็นผลเนื่องมาจากการระมัดระวังมากเกินไป หรือเกิดจากการเข้าใจผิด ตัวอย่างเหตุการณ์ได้แก่ ตัวตรวจจับการบุกรุกแบบดอส (DoS : Denial-of-Service) ได้พบเหตุการณ์ที่น่าสงสัยและหลังจากการวิเคราะห์ก็สรุปว่าเหตุการณ์ดังกล่าวเป็นการบุกรุกแบบหนึ่ง ทำให้ระบบจำเป็นต้องหยุดให้บริการนั้นไปชั่วคราว ซึ่งส่งผลทำให้ผู้ที่จะขอเข้าใช้บริการคนอื่นไม่สามารถทำงานได้ตามปกติ หลังจากได้ตรวจสอบแน่ชัดแล้วสรุปว่ามีการตรวจจับที่ผิดพลาด จากเหตุการณ์ดังกล่าวเห็นได้ชัดว่าผลเสียที่เกิดขึ้นเป็นผลทำให้ผู้ใช้งานปกติไม่สามารถขอเข้าใช้บริการได้ตามปกติได้ การตัดสินใจผิดทางบวกแสดงในรูปที่ 3

การตัดสินใจผิดทางลบจะส่งผลเสียมากกว่าแบบแรก นั่นคือหากมีเหตุการณ์ที่ผิดปกติเกิดขึ้น แต่ตัวตรวจจับการบุกรุกสรุปว่าเป็นเหตุการณ์ปกติ ซึ่งผลที่เกิดขึ้นคือผู้บุกรุกสามารถเข้าสู่ระบบและสร้างความเสียหายได้โดยง่าย เนื่องจากไม่มีระบบระวังภัยอีกต่อไป การตัดสินใจผิดทางลบจะเกิดขึ้นเนื่องจากความสามารถของการวิเคราะห์เหตุการณ์ที่ไม่มีประสิทธิภาพพอ หรืออาจจะเกิดจากความตั้งใจของผู้บุกรุกเอง วิธีการที่ผู้ไม่ประสงค์ดีนิยมทำกันมากในปัจจุบันคือ การทำให้ตัวตรวจจับการบุกรุกไม่สามารถทำงานได้ตามปกติซึ่งเป็นผลทำให้ไม่มีระบบเฝ้าระวัง หากจะเปรียบเทียบระบบคอมพิวเตอร์เป็นร่างกายมนุษย์แล้ว การทำให้ตัวตรวจจับการบุกรุกทำงานไม่ได้ก็หมายถึงการทำให้ภูมิคุ้มกันของร่างกายบกพร่องหรือระบบป้องกันต่าง ๆ ของร่างกายไม่สามารถทำงานได้นั่นเอง ซึ่งจะเป็นผลเสียอย่างยิ่ง การตัดสินใจผิดทางลบแสดงในรูปที่ 4

ตัวตรวจจับการบุกรุกที่ทำงานแบบเวลาจริง

เมื่อตัวตรวจจับการบุกรุกวิเคราะห์เหตุการณ์ที่ผิดปกติแล้วรายงานความผิดพลาดที่เกิดขึ้นให้ผู้ดูแลระบบทราบ ยิ่งถ้าหากมีการวิเคราะห์ได้รวดเร็วมากเท่าใดก็จะยิ่งทำให้สามารถจัดการกับปัญหานั้นได้ดียิ่งขึ้น ดังนั้นตัวตรวจจับการบุกรุกที่ดีจึงควรจะมีการทำงานที่รวดเร็ว เป็นที่ทราบว่าในปัจจุบันยังไม่มีตัวตรวจจับการบุกรุกระบบใดที่สามารถทำงานเป็นแบบเวลาจริงได้ การทำงานแบบเวลาจริงคือการทำงานของระบบที่สามารถตอบโต้กับเหตุการณ์ได้ทันท่วงทีหรือทันทีทันใด เช่น หากมีการบุกรุกเกิดขึ้นก็สามารถวิเคราะห์ได้ในขณะที่เหตุการณ์นั้นกำลังดำเนินอยู่และสามารถตอบโต้ได้อย่างทันทีโดยไม่เสียเวลา โดยปกติแล้วระบบที่เป็นแบบเวลาจริงนั้นจะถือเป็นระบบที่อยู่ในอุดมคติ เพราะในปัจจุบันยังมีข้อจำกัดมากมายที่ทำให้ไม่สามารถทำได้

อุปสรรคสำคัญที่ทำให้ระบบไม่สามารถทำงานเป็นแบบเวลาจริงได้คือ ประการแรก เป็นเหตุผลเนื่องมาจากข้อจำกัดของอุปกรณ์ทางด้านฮาร์ดแวร์ของระบบ เช่น ไม่สามารถเข้าถึงหรือดึงข้อมูลจากดิสก์ได้ในเวลาที่รวดเร็วเพียงพอ หรือซีพียูไม่สามารถที่จะวิเคราะห์ข้อมูลได้อย่างเร็วพอ ประการที่สองเกิดขึ้นเนื่องจากความจำเป็นที่จะต้องวิเคราะห์ข้อมูลจำนวนมากซึ่งเป็นผลทำให้ต้องใช้เวลามากตามไปด้วย เพราะในบางกรณีมีความจำเป็นจะต้องพิจารณาข้อมูลเพิ่มเพื่อที่จะสรุปผลการวิเคราะห์ให้ได้ ประการสุดท้ายเกิดขึ้นเนื่องจากโปรแกรมที่ใช้งานไม่มีประสิทธิภาพเพียงพอในแง่ของระเบียบขั้นตอนวิธีหรืออัลกอริธึมไม่มีประสิทธิภาพนั่นเอง

บทสรุป

จากรายละเอียดต่าง ๆ ข้างต้นเป็นการอธิบายถึงปัญหาสำคัญที่ทำให้ระบบตรวจจับการบุกรุกทำงานไม่ได้ดีเท่าที่ควร ซึ่งจะเห็นได้ชัดว่าปัญหาต่าง ๆ เหล่านี้ล้วนเกี่ยวเนื่องกันไปทั้งหมด สิ่งที่นักวิจัยกำลังให้ความสนใจกันมากในปัจจุบันคือ การคิดค้นหาวิธีการเพื่อลดข้อจำกัดและปัญหาหลัก ๆ ของตัวตรวจจับการบุกรุกให้เหลือน้อยที่สุด ซึ่งเมื่อถึงตอนนั้นก็จะมีระบบตรวจจับการบุกรุกที่มีประสิทธิภาพดียิ่งขึ้นให้ได้ใช้งานกันต่อไป